An Analysis of Gear Fault Detection Methods as Applied to Pitting Fatigue Failure Data

/N -37

/6 7252

P- 12

J.J. Zakrajsek and D.P. Townsend
Lewis Research Center
Cleveland, Ohio

H.J. Decker
Propulsion Directorate
U.S. Army Aviation Systems Command
Lewis Research Center
Cleveland, Ohio

NASA

US ARMY
AVIATION
SYSTEMS COMMAND

AN ANALYSIS OF GEAR FAULT DETECTION METHODS AS APPLIED TO PITTING FATIGUE FAILURE DATA

James J. Zakrajsek and Dennis P. Townsend
National Aeronautics and Space Administration
Lewis Research Center
21000 Brookpark Road
Cleveland, Ohio 44135

Harry J. Decker
Propulsion Directorate
U.S. Army Aviation Systems Command
Lewis Research Center
21000 Brookpark Road
Cleveland, Ohio 44135

Abstract: The application of gear fault prediction techniques to experimental data is examined. A single mesh spur gear fatigue rig was used to produce naturally occurring faults on a number of test gear sets. Gear tooth surface pitting was the primary failure mode for a majority of the test runs. The damage ranged from moderate pitting on two teeth in one test to spalling on several teeth in another test. Previously published failure prediction techniques were applied to the data as it was acquired to provide a means of monitoring the test and stopping it when a failure was suspected. A newly developed technique along with variations of published methods were also applied to the experimental data. The published methods experienced some success in detecting initial pitting before it progressed to affect the overall root-mean-square (RMS) vibration level. The new technique robustly detected the damage on all of the tests, and in most cases continued to react to the damage as it spread and increased in severity. Since no single method was able to consistently predict the damage first on all the runs, it was concluded that the best approach to reliably detect pitting damage is to use a combination of detection methods.

Key Words: Gear; Fatigue; Diagnostics; Failure Prediction

Introduction: Drive train diagnostics is becoming one of the most significant areas of research in rotorcraft propulsion. The need for a reliable health and usage monitoring system for the propulsion system can be seen by reviewing some rotorcraft accident statistics. An investigation of serious rotorcraft accidents that were a result of fatigue failures showed that 32 percent were due to engine and transmission components [1]. Also, 60 percent of the serious rotorcraft accidents were found to occur during cruise flight. Civil helicopters need a thirtyfold increase in their safety record to equal that of conventional fixed-wing turbojet aircraft. Practically, this can only be accomplished with the aid of a highly reliable, on-line health and usage monitoring unit. Diagnostic research is required to develop and prove various fault detection concepts and methodologies.

A number of methods have been developed to provide early detection of gear tooth surface damage. McFadden proposed a method to detect gear tooth cracks and spalls by demodulating the time signal [3]. Stewart devised several time domain discriminant methods

of which FM0, a coarse fault detection parameter, and FM4, an isolated fault detection parameter, are the most widely referenced [4]. Martin proposed using the sixth and eighth statistical moments of the time signal to detect surface damage [2]. A new method, NA4, was recently developed at NASA Lewis Research Center to detect and continue to react to gear tooth surface damage as it spreads and grows in severity.

Verification of these detection methods with experimental data along with a comparison of their relative performance is a crucial step in the overall process of developing a highly reliable health monitoring system.

In view of the aforementioned, it becomes the object of the research reported herein to determine the relative performance of the detection methods as they are applied to experimental data. Each method is applied to vibration data obtained from a gear fatigue test rig at NASA Lewis, where test gears are run until a fatigue failure occurs. The failure modes of the five tests used in this study ranged from moderate pitting on two teeth in one test to spalling on several teeth in another test. Results of each method are compared for each test, and overall conclusions are made regarding the performance of the methods.

<u>Theory of Fault Detection Methods</u>: All of the methods in this investigation utilized vibration data that was preprocessed as it was collected. To eliminate the noise and vibration that is incoherent with the rotational speed of the test gears, the raw vibration data was time synchronous averaged immediately after being digitized. During time synchronous averaging, the data was also interpolated to obtain 1024 points per revolution of the test gears. Each of the methods presented below were then applied to the time averaged and interpolated vibration data.

FM0 is formulated to be a robust indicator of major faults in a gear mesh by detecting major changes in the meshing pattern [4]. FM0 is found by dividing the peak-to-peak level of the signal average by the sum of the amplitudes of the mesh frequency and its harmonics. In major tooth faults, such as breakage, the peak-to-peak level tends to increase, resulting in FM0 increasing. For heavy distributed wear or damage, the peak-to-peak remains somewhat constant but the meshing frequency levels tend to decrease, resulting in FM0 increasing.

FM4 was developed to detect changes in the vibration pattern resulting from damage on a limited number of teeth [4]. A difference signal is first constructed by removing the regular meshing components (shaft frequency and harmonics, primary meshing frequency and harmonics along with their first order sidebands) from the original signal. The fourth normalized statistical moment (normalized kurtosis) is then applied to this difference signal. For a gear in good condition the difference signal would be primarily noise with a Gaussian amplitude distribution, resulting in a normalized kurtosis value of 3 (nondimensional). When one or two teeth develop a defect (such as a crack, pit, or spall) a peak or series of peaks appear in the difference signal, resulting in the normalized kurtosis value to increase beyond the nominal value of 3.

A demodulation technique was developed to detect local gear defects such as fatigue cracks, pits and spalls [3]. The basic theory behind this technique is that a gear tooth defect will produce sidebands that modulate the dominant meshing frequency. In this method, the signal is band-passed filtered about a dominant meshing frequency, including as many sidebands as possible. The Hilbert transform is then used to convert the real band-passed signal into a complex time signal, or analytic signal. The normalized kurtosis is then applied to the amplitude modulation function (magnitude of the analytic signal) in an attempt to

detect gear tooth damage through the modulating sidebands. Again, a value of 3 would indicate a nominal condition, and a value over 3 indicates possible tooth damage.

M6A and M8A are variations of the sixth (M6) and eighth (M8) normalized statistical moments proposed by Martin to detect surface damage using vibration signals [2]. M6 and M8 are applied to the same difference signal as defined in the definition of FM4. The basic theory behind M6A and M8A is the same as that for FM4, except M6A and M8A should be more sensitive to peaks in the difference signal. Also, the value for nominal conditions (Gaussian distribution) is 15 for M6A, and 105 for M8A.

NA4 is a new method that was developed by the authors to not only detect the onset of damage, as FM4 does, but also to continue to react to the damage as it spreads and increases in magnitude. Similar to FM4, a residual signal is constructed by removing regular meshing components from the original signal, however, for NA4, the first order sidebands stay in the residual signal. The fourth statistical moment of the residual signal is then divided by the current run time averaged variance of the residual signal, raised to the second power, resulting in the quasi-normalized kurtosis given below:

$$
NA4(M) = \frac{N \sum\limits_{i=1}^{N} (r_i - \bar{r})^4}{\left\{ \dfrac{1}{M} \sum\limits_{j=1}^{M} \left[\sum\limits_{i=1}^{N} (r_{ij} - \bar{r})^2 \right] \right\}^2}
$$

where

r	residual signal
$\bar{r}$	mean value of residual signal
N	total number of data points in time record
i	data point number in time record
M	current time record number in run ensemble
j	time record number in run ensemble

In NA4, the kurtosis is normalized, however it is normalized using the variance of the residual signal averaged over the run up to point in the run that NA4 is being calculated for. With this method, the changes in the residual signal are constantly being compared to the running average of the variance of the system, or a weighted baseline for the specific system in "good" condition. This should allow NA4 to grow with the severity of the fault until the average of the variance itself changes. As with FM4, NA4 is dimensionless, with a value of 3 under nominal conditions.

Apparatus and Gear Damage Review: A spur gear fatigue rig at NASA Lewis was used to obtain experimental data. The primary purpose of this rig is to study the effects of gear materials, gear surface treatments, and lubrication types on the surface fatigue strength of aircraft quality gears. The rig was recently modified to allow it to be used for diagnostic studies as well as fatigue research [5]. Vibration data from an accelerometer mounted on a bearing end plate was captured using an on-line program running on a personal computer with an analog to digital conversion board and anti-aliasing filter. The test gears are standard spur gears having 28 teeth and a pitch diameter of 88.9 mm (3.50 in.). The gears were loaded to 74.6 Nm (660 in.-lb) at an operating speed of 10,000 rpm.

Some examples of the different magnitudes of tooth damage found in the five tests (runs 1 to 5) of this study are illustrated in Figure 1. Figure 1(a) shows the isolated heavy pitting damage that was found on the test gears in run 1 at 131 hours into the test. Figure 1(b) shows an example of the spalling damage found at the end of the test of run 1. Figure 1(c) illustrates an example of the moderate pitting found in the tests. Similarly, Figure 1(d) gives an example of the heavy pitting damage found in the tests. Details of the damage found in each test are given below, with Figure 1 serving as a pictorial reference of damage magnitude.

At 131 hours into run 1, damage was found on two teeth on the driver gear (one heavy and one moderate pitting). Both mating teeth on the driven gear were also found to be damaged (both heavy pitting). Figure 1(a) illustrates the heavy pitting damage on the driver and driven gears at 131 hours. At the end of run 1, spalling (Figure 1(b)) and heavy pitting damage was found on roughly one third of the teeth on both the driver and driven gears.

At the end of run 2, damage was found on three consecutive teeth on the driver gear (one heavy and two moderate pitting). Two of the three mating teeth on the driven gear were also found to be damaged (both moderate pitting).

At the end of run 3, damage was found on four consecutive teeth on the driver gear (one spalling, two heavy, and one moderate pitting). One of the four mating teeth on the driven gear was also found to be damaged (moderate pitting).

At the end of run 4, damage was found on two consecutive teeth on the driver gear (both heavy pitting). The two mating teeth on the driven gear were also found to be damaged (one heavy, and one moderate pitting).

At 294 hours into run 5, micropitting and wear was found on nearly all the teeth of the driver gear. At the end of run 5, moderate pitting was found on eight teeth distributed on the driver gear. Three consecutive teeth on the driven gear were found to have moderate pitting damage.

Discussion of Results: The results of applying the fault detection methods to the experimentally obtained vibration data are illustrated in Figures 2 to 6.

Figure 2 presents the results of all the parameters for run 1. The vertical centerline in each plot represents the point in time ($t = 131$ hours) in which the rig was stopped and the damage was recorded, as shown in Figure 1(a). As seen in Figure 2, the parameters FM4, NA4, Kurtosis of AMF, M6A, and M8A all detect tooth damage at $t = 110$ hours, or 25 hours before FM0 reacts, and 27 hours before the overall root-mean-square (RMS) vibration level increases. FM4 peaked at a value of 5.4, then dropped off to the nominal value of 3 at $t = 131$ hours. It is possible that only one of the two teeth found damaged at $t = 131$ hours actually started at the time FM4 reacted, and when the damage spread to the other tooth, FM4 lost its sensitivity by decreasing back to its nominal value. The results of the demodulation method for run 1 (Figure 2(e)), are the best results obtained from that method. In other runs it showed results very similar to FM4 results (runs 2 and 3), or gave no indication at all (runs 4 and 5). As seen in Figure 2, M6A and M8A results follow the same trends indicated by FM4. M6A and M8A, however reacted more strongly to the damage, as indicated by the 300 percent and 863 percent increases over nominal values for M6A and M8A, respectively, as compared to an 80 percent increase for FM4. These results for M6A and M8A are very typical of the results obtained for M6A and M8A on the other four runs. FM0 gave a solid indication of over three times its nominal value, and 2 hours in

advance of the RMS level increase. NA4 gave the best performance for run 1. Figure 2(f) shows the first 135 hours of Figure 2(d), with an expanded vertical scale, for clarity. As seen in these two figures, NA4 reacts very robustly to the start of damage, sharply increasing to a value of 25, and remains somewhat steady at a value of 15 even as the other parameters (FM4, M6A, etc.) drop back down to nominal values. NA4 then increases sharply to a peak value of 230, following a trend similar to the RMS level increase. This could be the point at which the extremely heavy damage started (as seen in Figure 1(b)), continuing to the end of the run.

As seen in Figure 3, the parameters FM0, FM4, and NA4 all react sharply to the tooth damage at 94 hours into run 2. FM0 reacted robustly to the damage, increasing to over double its nominal value, whereas the overall RMS vibration level gradually increases with run time. FM4 also reacted by increasing from a value a little under the nominal 3 to a relatively steady value of 4.5 through to the end of the run. Because the heavy pitting damage was still isolated to only one of the three damaged teeth on the driver, FM4 was able to continue to react to the damage. NA4 gave the most robust reaction to the damage, increasing sharply from the nominal value of 3 to a value of 9 at t = 94 hours. NA4 then continues to increase from 9 to a peak of 29, growing gradually with the damage until 2 hours before the end where NA4 then drops off, due to a sharp increase in the denominator of NA4.

In run 3, only FM0 showed any significant reaction to the start of damage at 43 hours into the run, as seen in Figure 4. FM0 increased to over double its nominal value at this time. The damage may have been too subtle for the overall RMS level to increase, and may have started somewhat simultaneously over the four driver teeth for FM4 to indicate only a low grade response at t = 43 hours. When FM4 and NA4 do react at t = 74 hours, possibly due to the spalling formation on one of the four driver teeth, NA4 again reacts more robustly, increasing to 8, as compared to 5 for FM4. Both parameters increase, but FM4 peaks at 7.5, whereas NA4 peaks at 43.

As illustrated in Figure 5, the damage in run 4 was detected by FM0 and FM4 at the same time that the overall RMS vibration level increased. FM0 again shows a significant reaction to the damage, increasing in value to nearly three times its nominal value, as compared to the RMS level, which increases only 40 percent over its nominal value. FM4 peaks at 4.6, then proceeds to fall back to the nominal value. One of the two heavily damaged teeth on the driver gear may have started first, followed by heavy damage on the second tooth and the resulting decrease in the response and thus sensitivity of FM4. NA4 gives a strong indication of damage nearly 5 hours before the other parameters, and peaks at the value of 18.5, as compared to 4.6 for FM4. NA4 then decreases after the peak to 6.5, as its denominator increases, at the end of the run.

The vertical centerline in all the plots in Figure 6 indicate the point in time (t = 294 hours) that run 5 was stopped and micropitting was found on nearly all the teeth on the driver gear. As seen in Figure 6, FM0 and NA4 clearly detect the micropitting damage. After this point, FM0 and NA4 increase sharply, with FM0 peaking at over twice its nominal value, and NA4 increasing to a value of 15, then slowly dropping off. The sharp increase seen in FM0, NA4, and even the overall RMS vibration level most probably corresponds to the initiation of the moderate pitting found on a number of teeth on both driver and driven gears at the end of the run. As evident in Figure 6(b), FM4 gave no indication of either the initial micropitting damage nor the moderate pitting damage found at the end of the run. Due to the nature of the damage, both the micropitting and moderate pitting damage may

have occurred simultaneously on more than one or two isolated teeth, FM4 was incapable of reacting to it.

Based on the results just presented, it is clearly evident that of all the methods investigated in this study, the previously published method FM0 and the newly developed method NA4 are the most robust and reliable indicators of gear tooth pitting fatigue damage. FM0 gave a clear indication of the pitting fatigue damage on all five runs. On an average, FM0 increased to nearly three times its nominal value several hours before the RMS level showed any real change, on a majority of the runs. NA4 also gave a clear indication of the pitting fatigue damage on all five runs. NA4 reacted not only to isolated pitting damage on one or two teeth, but also to pitting damage that occurred over a number of teeth around the gear. NA4 gave robust initial reactions to the damage, increasing from the nominal value of 3 to an average value of 15, and in some cases continued to react as the damage spread and/or increased in severity.

The other methods were able to predict the pitting damage in most of the runs, however, they did not perform as reliably or robustly as FM0 and NA4. FM4 is a relatively good indicator of damage on one or two isolated teeth, however, results showed that as the damage spread to other teeth FM4 lost its sensitivity and dropped back down to the nominal value of 3. In one case FM4 never reacted, as the damage may have initiated on a number of teeth at approximately the same time. Although M6A and M8A showed stronger reactions to the damage, as compared to FM4, they exhibited the same trends as FM4, and thus the same weaknesses. The demodulation method gave results no better than FM4 in three of the runs, and failed to react to the damage in the remaining two runs.

In order to accurately and reliably detect gear tooth pitting fatigue damage, several methods need to be used in combination. Even with the limited data used in this study, not one method was able to give a first indication of the damage consistently on all five runs. Several methods, FM0 and NA4 as a minimum, need to operate in parallel in order to provide a reliable way of detecting the pitting damage as far in advance of severe damage as possible.

Conclusions: Based on this investigation, the following conclusions can be made

1) The newly developed parameter, NA4, reacted very robustly to the damage on all the runs. It reacted to isolated pitting damage as well as pitting damage on a number of teeth distributed around the gear. In several cases, NA4 continued to react as the damage spread and/or increased in severity, thus indicating damage level.

2) FM0 is a strong indicator of gear tooth pitting damage occurring over a number of teeth on a gear. For a majority of the runs, FM0 reacted to the damage before the RMS vibration level reacted. On those runs where FM0 reacted at the same time as the RMS level, FM0 gave a much clearer indication.

3) FM4 reacts well to damage on one or two isolated teeth, but loses its sensitivity significantly as the damage spreads to other teeth. FM4 failed to detect damage on one run as the pitting damage may have initiated on several teeth at the same period in time.

4) M6A and M8A exhibited stronger reactions to the damage, as compared to FM4. They, however, showed the same trends, and thus the same weaknesses, as FM4.

5) No single method was able to consistently predict the pitting damage before the others on all the runs. A number of the methods, FM0 and NA4 as a minimum, need to be used in combination in order to reliably detect gear tooth pitting damage.

REFERENCES

1. Astridge, D.G.: Helicopter Transmissions—Design for Safety and Reliability. Inst. Mech. Eng. Proc., Pt. G—J. Aerosp. Eng., vol. 203, no. G2, 1989, pp. 123–138.

2. Martin, H.R.: Statistical Moment Analysis As a Means of Surface Damage Detection. Proceedings of the 7th International Modal Analysis Conference, Society for Experimental Mechanics, Schenectady, NY, 1989, pp. 1016–1021.

3. McFadden, P.D.: Detecting Fatigue Cracks in Gears by Amplitude and Phase Demodulation of the Meshing Vibration. J. Vib. Acoust. Stress Reliab. Design, vol. 108, no. 2, Apr. 1986, pp. 165–170.

4. Stewart, R.M.: Some Useful Data Analysis Techniques for Gearbox Diagnostics. Report MHM/R/10/77, Machine Health Monitoring Group, Institute of Sound and Vibration Research, University of Southampton, July 1977.

5. Zakrajsek, J.J., et al.: Analysis and Modification of a Single-Mesh Gear Fatigue Rig for Use in Diagnostic Studies. NASA TM–105416, 1992.
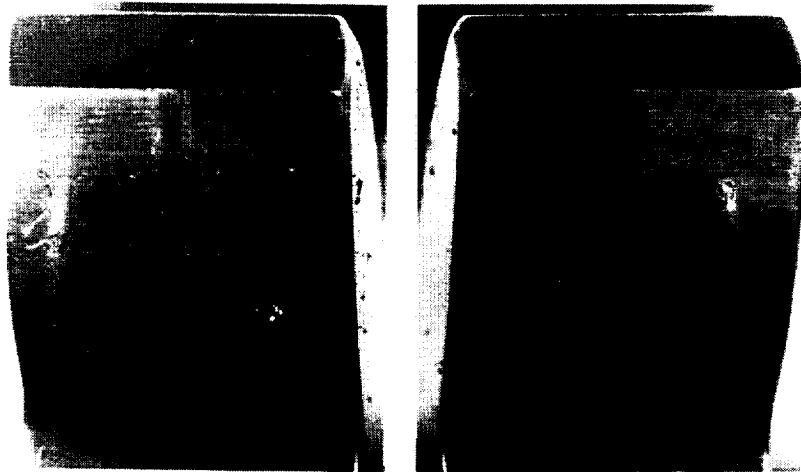
(a) Heavy pitting on two teeth in Run 1 at t = 131 hr into run.
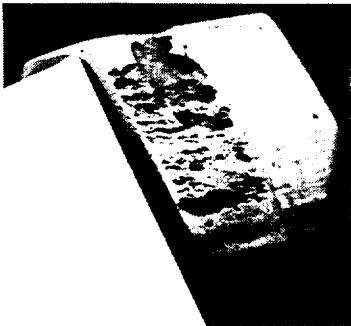
(b) Example of spalling on tooth in Run 1 at end of test (t = 198 hr).

(c) Example of moderate pitting (Run 5, end of test).

(d) Example of heavy pitting (Run 3, end of test).

Figure 1.—Examples of actual damage on gear teeth.
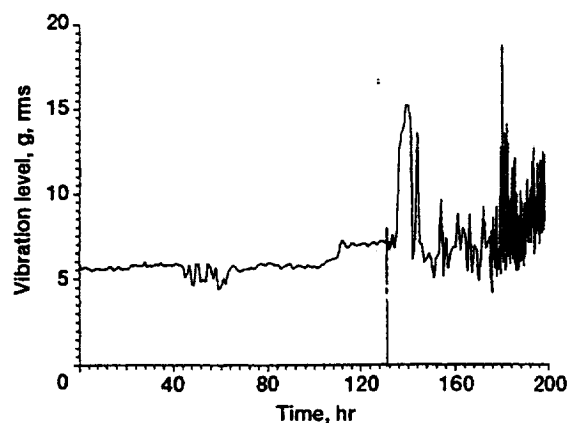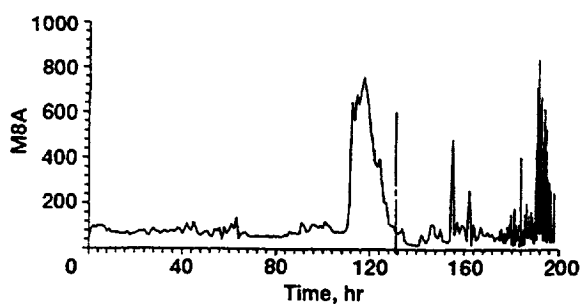
(a) RMS.
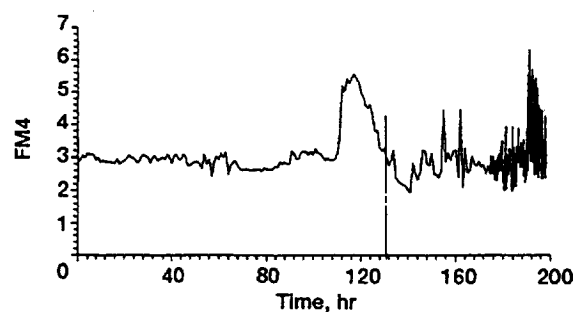
(b) FM4.
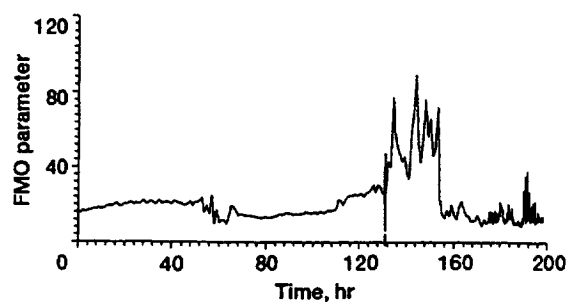
(c) FMO.
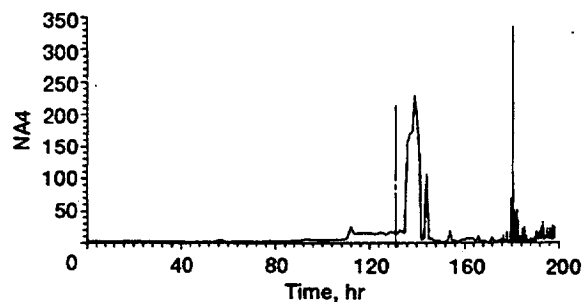
(d) NA4.
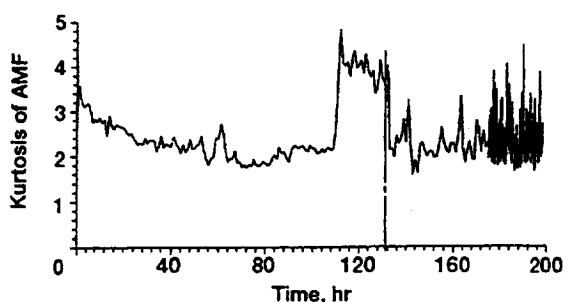
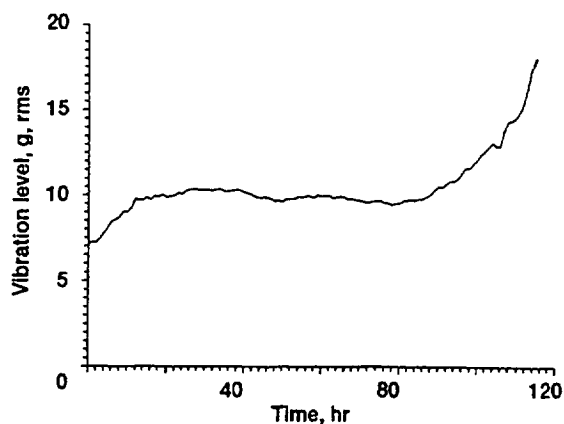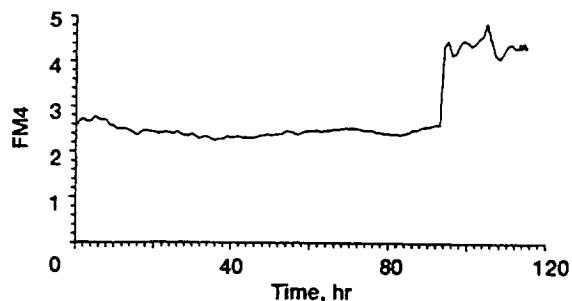(e) Kurtosis of amplitude modulation function.
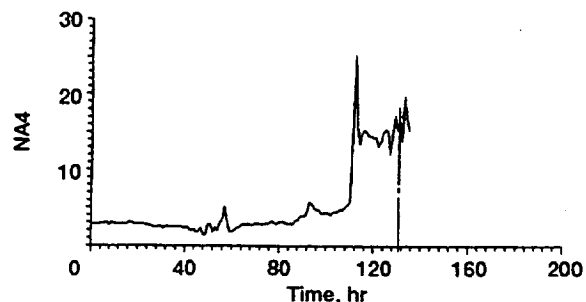
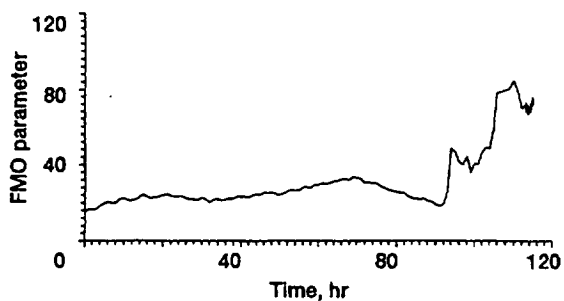(f) NA4 up to t = 135 hr..
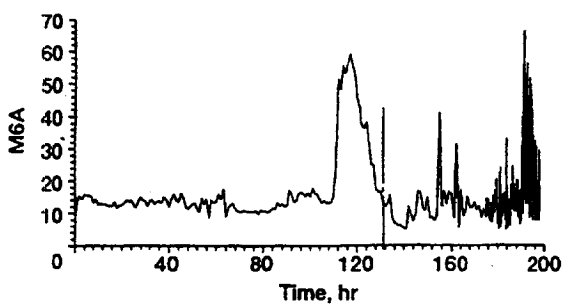
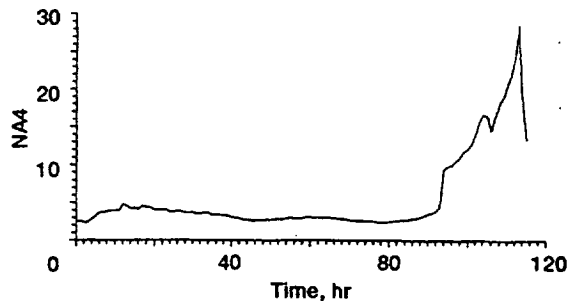(g) M6A.

(h) M8A.

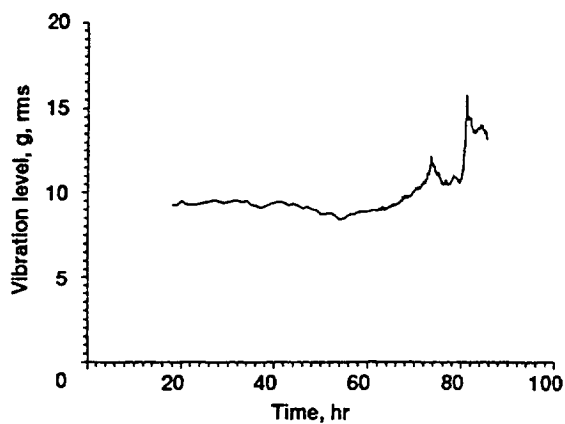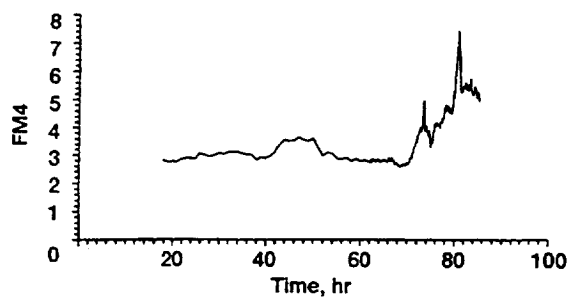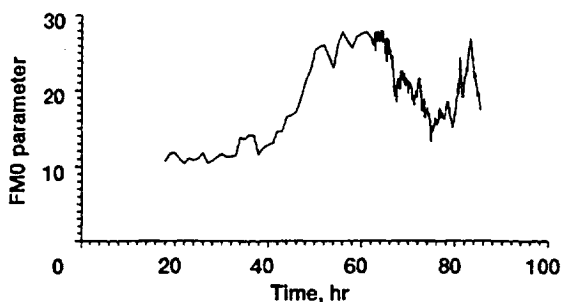Figure 2.—Run 1 results.

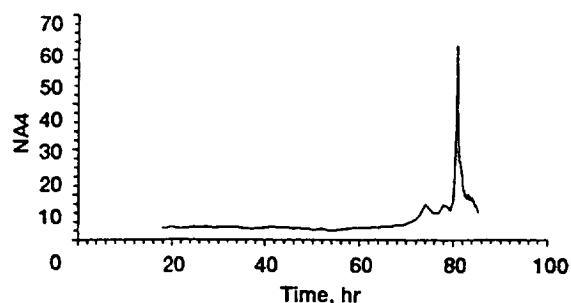(a) RMS level.

(b) FM4.

(c) FMO.

(d) NA4.

Figure 3.—Run 2 results.

(a) RMS level.

(b) FM4.

(c) FM0.

(d) NA4.

Figure 4.—Run 3 results.

9

(a) RMS level.

(b) FM4.

(c) FMO.

(d) NA4.

Figure 5.—Run 4 results.

(a) RMS level.

(b) FM4.

(c) FMO.

(d) NA4.
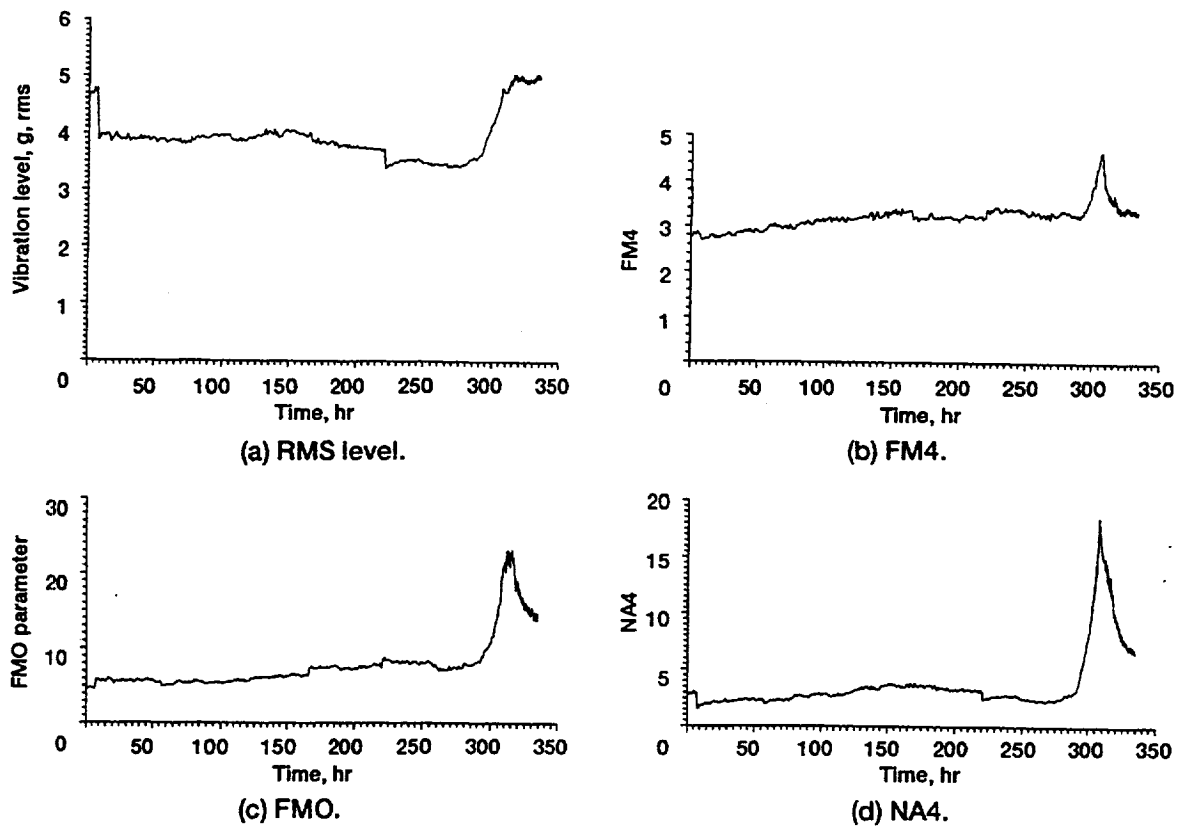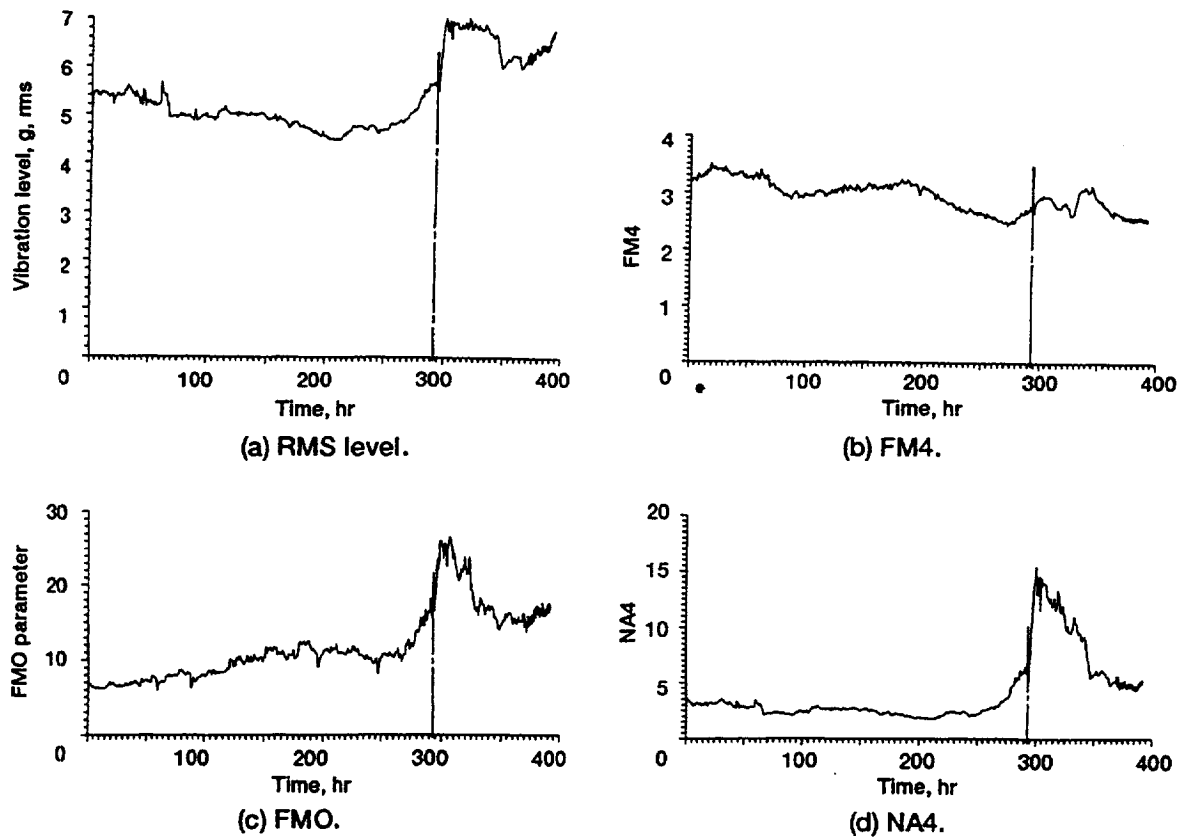
Figure 6.—Run 5 results.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 1993	Technical Memorandum

4. TITLE AND SUBTITLE

An Analysis of Gear Fault Detection Methods as Applied to Pitting Fatigue Failure Data

5. FUNDING NUMBERS

WU–505–62–10

6. AUTHOR(S)

J.J. Zakrajsek, D.P. Townsend, and H.J. Decker

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Lewis Research Center
Cleveland, Ohio 44135–3191
and
Propulsion Directorate
U.S. Army Aviation Systems Command
Cleveland, Ohio 44135–3191

8. PERFORMING ORGANIZATION REPORT NUMBER

E–7470

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, D.C. 20546–0001
and
U.S. Army Aviation Systems Command
St. Louis, Mo. 63120–1798

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA TM–105950
AVSCOM 92–C–035

11. SUPPLEMENTARY NOTES

Prepared for the 47th Mechanical Failure Prevention Group, Virginia Beach, Virginia, April 13–15, 1993. J.J. Zakrajsek and D.P. Townsend, NASA Lewis Research Center. H.J. Decker, Propulsion Directorate, U.S. Army Aviation Systems Command. Responsible person. J.J. Zakrajsek, (216) 433–3968.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category 37

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The application of gear fault prediction techniques to experimental data is examined. A single mesh spur gear fatigue rig was used to produce naturally occurring faults on a number of test gear sets. Gear tooth surface pitting was the primary failure mode for a majority of the test runs. The damage ranged from moderate pitting on two teeth in one test to spalling on several teeth in another test. Previously published failure prediction techniques were applied to the data as it was acquired to provide a means of monitoring the test and stopping it when a failure was suspected. A newly developed technique along with variations of published methods were also applied to the experimental data. The published methods experienced some success in detecting initial pitting before it progressed to affect the overall root-mean-square (RMS) vibration level. The new technique robustly detected the damage on all of the tests, and in most cases continued to react to the damage as it spread and increased in severity. Since no single method was able to consistently predict the damage first on all the runs, it was concluded that the best approach to reliably detect pitting damage is to use a combination of detection methods.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Gear; Fatigue; Diagnostics; Failure Prediction	12
	16. PRICE CODE A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

NASA